A Chemoenzymatic Approach Toward the Rapid and Sensitive Detection of *O*-GlcNAc Posttranslational Modifications

- Supporting Information -

Nelly Khidekel,[†] Sabine Arndt,[†] Nathan Lamarre-Vincent,[†] Alexander Lippert,[†] Katherine G. Poulin-Kerstien,[†] Boopathy Ramakrishnan,[‡§] Pradman K. Qasba,[‡] and Linda C. Hsieh-Wilson[†*]

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, and Structural Glycobiology Section, Laboratory of Experimental and Computational Biology, CCR, NCI, SAIC-Frederick, Inc, Frederick, MD 21702-1201, USA.

General Methods:

Chemicals and reagents were used without further purification unless otherwise noted. If necessary, reactions were performed under argon atmosphere using anhydrous solvents. Thin layer chromatography was performed using E. Merck silica gel 60 F254 precoated plates and visualized using cerium ammonium molybdate stain. Flash column chromatography was carried out with Silica Gel 60 (230-400 mesh). NMR spectra were obtained on a Varian Mercury 300 instrument. High resolution mass spectra were obtained with a Jeol JMS-600H spectrometer. The peptide TAPTS(*O*-GlcNAc)TIAPG was synthesized at the Beckman Institute Biopolymer Synthesis Center using standard Fmoc chemistry. The Fmoc-protected, peracetylated *O*-GlcNAc serine amino acid was kindly synthesized by S. Tully as reported by Seitz *et al.*[1] Baculovirus preparation and protein expression of CREB in *Spodoptera frugiperda* (Sf9) cells were performed by Dr. P. Snow at the Beckman Institute Protein Expression Facility at the California Institute of Technology.[2] HeLa cell nuclear extracts were kindly prepared in H.-C. Tai according to published procedures.[3] Y289L and wild-type GalT were expressed and purified as described previously.[4] All protein concentrations were measured using the Bradford assay (Bio-Rad Laboratories, Hercules, CA).

General Reagents:

Unless otherwise noted, reagents were purchased from the commercial suppliers Fisher (Fairlawn, NJ) and Sigma-Aldrich (St. Louis, MO) and were used without further purification. Protease inhibitors were purchased from Sigma-Aldrich or Alexis Biochemicals (San Diego, CA). Bovine GalT, ovalbumin, and α-crystallin were obtained from Sigma-Aldrich. Uridine diphospho-D-[6-^{3}H]galactose, Hyperfilm ECL and Amplify reagent were purchased from Amersham Biosciences (Piscataway, NJ). WGA lectin was purchased from E-Y Laboratories (San Mateo, CA). RL-2 antibody was purchased from Affinity Bioreagents (Golden, CO). Alkaline phosphatase was purchased from New England Biolabs (Beverly, MA), and bovine serum albumin (BSA) was obtained from Fisher. SuperSignal West Pico chemiluminescence reagents and secondary antibodies were from Pierce (Rockford, IL), and the CTD110.6 antibody was from Covance Research Products (Berkeley, CA). Nitrocellulose was from Schleicher and Schuell (Keene, NH), and PVDF was from Millipore (Bedford, MA).

2-Acetonyl-2-deoxy-3,4,5-tri-O-acetyl-β-D-galactopyranose[5]

Ketone **2** (289 mg, 0.744 mmol) was dissolved in acetonitrile (1.5 mL), and Me_2NH in THF (2.0 M solution, 2.80 mL, 5.60 mmol) was added. The reaction mixture was stirred for 24 h at r.t.

The solvents and reagents were evaporated *in vacuo*. Flash chromatography on silica gel (1:1 hexanes: EtOAc) gave the monodeacetylated product (136 mg, 0.393 mmol, 53%) as a colorless oil.
^{1}H NMR (300 MHz, $CDCl_3$): δ 5.49-5.46 (m, 1H, 1-H), 5.34-5.33 (m, 1H, 4-H), 5.10 (dd, *J* = 12.0, 3.0 Hz, 1H, 3-H), 4.39 (t, *J* = 6.6 Hz, 1H, 5-H), 4.18-4.04 (m, 2H, 6-H_2), 2.84-2.72 (m, 1H, 2-H), 2.62-2.54 (m, 2H, 1'-H_2), 2.17, 2.14, 2.06, 2.01 (4 x s, 12H, 3 x Ac, 3'-H_3).
^{13}C NMR (75 MHz, $CDCl_3$): δ 207.1, 170.4, 170.3, 170.2, 92.8, 68.7, 66.7, 66.1, 62.3, 40.9, 34.7, 30.4, 20.7, 20.7, 20.7.
HRMS(FAB) calcd. for $C_{15}H_{23}O_9$ $[M+H]^+$ 347.1342, found 347.1342.

Dibenzyl (2-acetonyl-2-deoxy-3,4,5-tri-O-acetyl-α-D-galactopyranosyl) phophate (3) [6]
The deprotected ketone (90 mg, 0.26 mmol) and 1*H*-tetrazole (91 mg, 1.3 mmol) were dissolved in dichloromethane (3 mL). The reaction mixture was cooled to -30 °C and dibenzyl *N,N'*-diisopropylphosphamidite (170 µL, 0.52 mmol) was added. The reaction mixture was warmed to r.t. over 30 min and stirred at r.t. After 1 h, the reaction mixture was again cooled to - 30 °C, and mCPBA (229 mg, 1.30 mmol) was added. The mixture was then stirred at 0 °C for 1 h and at r.t. for 1 h. The reaction was subsequently diluted in dichloromethane, washed twice with 10% Na_2SO_3, once with $NaHCO_3$, and once with H_2O. The organic phase was dried over $MgSO_4$, filtered and concentrated. Flash chromatography on silica gel (1:1 hexanes:EtOAc) gave **3** (83 mg, 0.14 mmol, 54%) as a colorless oil.
^{1}H NMR (300 MHz, $CDCl_3$): δ 7.34-7.32 (m, 10H, arom), 5.86 (dd, *J* = 6.0, 3.3 Hz, 1H, 1-H), 5.29 (m, 1H, 4-H), 5.15-4.98 (m, 4H, bn), 4.92 (dd, *J* = 2.7, 12.0 Hz, 1H, 3-H), 4.25 (t, *J* = 6.5 Hz, 1H, 5-H), 4.07-3.93 (m, 2H, 6-H_2), 2.90-2.80 (m, 1H, 2-H), 2.35 (d, *J* = 7.2 Hz, 2H, 1'-H_2), 2.09, 1.95, 1.91, 1.87 (4 x s, 12H, 3 x ac, 3'-H_2).
^{31}P NMR (121 MHz, $CDCl_3$): δ -1.31.
^{13}C NMR (75 MHz, $CDCl_3$): δ 205.7, 170.0, 170.0, 169.8, 128.6, 128.5, 128.5, 127.9, 97.7(d), 69.6(d), 69.5, 68.3, 68.0, 65.9, 61.7, 39.7, 34.4(d), 29.9, 20.6, 20.6, 20.5.
HRMS(FAB): calcd. for $C_{29}H_{36}O_{12}P$ $[M+H]^+$ 607.1945, found 607.1924.

Uridine 5'-diphospho-2-acetonyl-2deoxy-α-D-galactopyranose diammonium salt (1) [7]
A solution of dibenzyl phosphate **3** (80 mg, 0.13 mmol) and tri-*n*-octylamine (35 µL) in methanol (10 mL) was hydrogenolyzed in the presence of 10% Pd/C (100 mg) under 1 atm H_2 for 20 h. The mixture was filtered, concentrated, dried and directly used in the next step.
UMP-morpholidate 4-morpholine-*N,N'*-dicyclohexylcarboxamidine salt (36 mg, 0.198 mmol) was added and the mixture was evaporated three times from anhydrous pyridine (1.5 mL). The mixture was dissolved in pyridine (1.0 mL), 1*H*-tetrazole (28 mg, 0.40 mmol) was added, and the solution was stirred for three days at r.t. After evaporation of the solvent, the reaction product was dissolved in a mixture of MeOH/water/TEA (2 mL/0.8 mL/0.4 mL) and stirred for 24 h. The residue was then dissolved in water and dichloromethane, and the organic phase was extracted twice with water. The aqueous phases were combined and lyophilized. The residue was purified on a Bio-Gel P2 (extra fine) column (1.5 x 80 cm), and eluted with 0.1 M NH_4HCO_3 at a flow rate of 0.6 mL/min. Lyophilization of the desired fractions (determined by HPLC, Varian Microsorb C18, 100 mM NH_4HCO_3, 4.1 min) gave **1** (38.7 mg, 0.060 mmol, 45%) as a colorless powder.

^{1}H NMR (300 MHz, D_2O): δ 7.96 (d, *J* = 8.1 Hz, 1H, 6''-H), 5.97-5.94 (m, 2H, 5''-H, 1'-H), 5.55 (dd, *J* = 7.8, 3.3 Hz, 1H, 1-H), 4.36-4.33 (m, 2H, 2'-H, 3'-H), 4.26-4.24 (m, 1H, 4'-H), 4.21-4.17 (m, 2H, 5'-H_2), 4.13 (t, *J* = 5.1 Hz, 1H, 5-H), 3.88 (m, 1H, 4-H), 3.79-3.69 (m, 3H, 3-H, 6-H_2), 2.79-2.75 (m, *J* = 4.2 Hz, 2H, 1'''-H_2), 2.53 (m, 1H, 2-H), 2.24 (s, 3H, 3'''-H_3).
^{31}P NMR (121 MHz, $CDCl_3$): δ -10.74 (d, *J* = 19.5 Hz), -12.06 (d, *J* = 20.1 Hz).
^{13}C NMR (75 MHz, D_2O): δ 214.3, 166.3, 151.9, 141.8, 102.9, 96.5, 88.6, 83.6, 74.0, 72.1, 69.9, 68.2, 65.1, 63.9, 61.6, 43.5, 41.6, 30.3.
HRMS(EI) calcd. for $C_{18}H_{27}O_{17}N_2P_2$ $[M-H]^-$ 605.0785, found 605.0803.

Labeling of the *O*-GlcNAc Peptide. The peptide TAPTS(*O*-GlcNAc)TIAPG (10μM) was dissolved in 25 mM MOPS buffer, pH 6.7 containing 5 mM $MnCl_2$ and 8 μM reference peptide (ThermoFinnigan, San Jose, CA). Ketone analogue **1** and mutant Y289L GalT were added to final concentrations of 1 mM and 100 ng/μL, respectively. Prior to enzyme addition, an aliquot of the reaction was removed as an initial time point for LC-MS analysis. Reactions were incubated at 4 °C for 6 h, after which an aliquot of the reaction mixture was removed for product analysis by LC-MS. The remainder of the reaction was diluted 5-fold into PBS (final concentration: 10.1 mM Na_2HPO_4, 1.76 mM KH_2HPO_4,, 137 mM NaCl, 2.7 mM KCl, pH 6.7), and *N*-(aminooxyacetyl)-*N'*-(D-biotinoyl) hydrazine (Molecular Probes, Eugene, OR) was added to a final concentration of 12 mM. After 8 h at 25 °C, the extent of biotin-oxime product was measured by LC-MS. Optimization of the experimental parameters suggested that a 6000:1 molar ratio of aminooxy biotin was optimal for complete conversion to the oxime product. Note that different batches of aminooxy biotin were found to contain variable amounts of TFA salts, affecting the final pH of the biotinylation reaction. Labeling reactions with wild-type GalT were performed identically, with the exception that reactions were incubated at 37 °C for 12 h.

LC-MS monitoring of *O*-GlcNAc peptide labeling reactions. Liquid chromatography and mass spectrometry (LC-MS) were performed on an LCQ Classic ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) interfaced with a Surveyor HPLC system (ThermoFinnigan, San Jose, CA). Approximately 10 pmoles of peptide from each labeling reaction was loaded onto a Luna column (2mm i.d. X 50mm) prepacked with 3μm 100Å C18 RP particles. Flow rate was maintained at 190 μL/min with a gradient optimized for separation of the *O*-GlcNAc peptide from labeled products. LC buffer A consisted of 2% CH_3CN in 0.1M aqueous AcOH and buffer B consisted of 90% CH_3CN in 0.1M aqueous AcOH. The gradient consisted of 0-3 min, 2% B; 3-6 min, 2-11%B; 11-14.5 min 11-27.5% B, 14.5-18min 27.5-100% B; 18-22min 100% B where the initial 5 minutes of flow were diverted to waste in order to avoid contamination of the mass spectrometer with salts. The LCQ was operated in automated mode using Xcalibur™ software. The electrospray voltage was 4.5 kV and the heated capillary was 200 °C. Ion injection time was set at 200 ms for full MS scan mode of operation (3 microscans per scan). The ion selection window was set at 500-1700 *m/z* for all experiments.

Supplementary Figure 1 shows the progress of the ketone labeling reaction using Y289L GalT and the subsequent reaction with aminooxy biotin, as monitored by LC-MS. Base peak chromatograms are shown before and 6 h after the addition of ketone analogue **1** and Y289L GalT. Complete conversion of the peptide to the desired ketone-labeled product was observed. For the aminooxy biotin reaction, formation of the oxime product was monitored using an extracted ion chromatogram within the mass range 1319.0-1321.0 *m/z* and 1633.0-1635.5 *m/z*, which was generated post-acquisition via the Xcalibur™ software. Extracted ion chromatograms

were necessary because the excess biotin in the reaction mixture dominated the base peak chromatograms. No appreciable amounts of the unbiotinylated starting material were observed after 8 h. Mass spectrometric analysis confirmed the identity of each product (Supplementary Figure 2).

Supplementary Figure 3 shows the progress of the labeling reaction using wild-type GalT after 12 h at 37 °C. The extent of conversion to ketone-labeled peptide was analyzed by measuring peak areas for the starting material (peak a) and product (peak b) using Xcalibur™ software, under the assumption that the *O*-GlcNAc peptide and its ketone-labeled analogue had similar ionization potentials. Approximately 1.5% of the desired product was formed with the wild-type GalT.

Labeling of CREB protein. Recombinant *O*-GlcNAc glycosylated CREB was generated by coexpression of CREB with *O*-GlcNAc glycosyltransferase in Sf9 cells as described previously.[2] 500 ng of CREB in 20 mM HEPES pH 7.9, 100 mM KCl, 0.2 mM EDTA, 15% glycerol was added to 50 mM MOPS pH 6.45 containing 5 mM $MnCl_2$ and 0.25 mU/μL alkaline phosphatase.[8] Analogue **1** and Y289L GalT were then added to final concentrations of 1 mM and 40 ng/μL, respectively. Control reactions without enzyme or analogue **1** were treated identically. Following incubation at 12 h at 4 °C, the reactions were diluted 5-fold into PBS containing protease inhibitors (5 μg/mL pepstatin, 5 μg/mL chymostatin, 20 μg/mL leupeptin, 20 μg/mL aprotinin, 20 μg/mL antipain, 0.2 mM PMSF). Aminooxy biotin was added to a final concentration of 2 mM, and the biotinylation reactions were incubated with gentle shaking for 12 h at 37 °C. Reactions were aliquoted for analysis and stopped by boiling in SDS-PAGE loading dye. Proteins were resolved by 10% SDS-PAGE, electrophoretically transferred to nitrocellulose, and probed with streptavidin-HRP.

Nitrocellulose blots were blocked for 1 h at RT using 3% periodated-BSA[9] in PBS, rinsed once with TBS (50 mM Tris•HCl, 150 mM NaCl, pH 7.4) containing 0.05% (v/v) tween-20, and probed with streptavidin-HRP (1:2500 to 1:5000) in TBS-0.05% tween for 1 h at RT. Note that we found some variability among different batches of streptavidin. In some cases, blots were probed for 1 h with streptavidin-HRP, rinsed several times with TBS-0.05% tween, and reprobed with another aliquot of streptavidin-HRP. After probing with streptavidin, membranes were rinsed and washed 5X10 min with TBS-0.1% tween containing 0.05% BSA. Streptavidin-HRP signal was visualized by chemiluminescence upon exposure to film. After streptavidin visualization, membranes were stripped in 5 mM Na_2HPO_4 pH 7.5, 2% SDS, and 2 mM βME, for 45 min at 60 °C, rinsed several times with dH_20, and re-probed with α-CREB antibody as previously described[2] with the modification that the antibody was used at a concentration of 1:400.

Labeling reactions with CREB expressed in *E. coli* were performed identically. To generate the bacterial protein, rat CREB cDNA was cloned into the prokaryotic expression vector pET23b(+) (Novagen, Madison, WI) using *HindIII* and *NdeI* restriction endonucleases. Electrocompetent BL21(DE3) cells were electroporated and grown in Luria-Bertani media supplemented with 100 mg/L ampicillin. Protein expression was induced with 0.3 mM isopropyl-β-D-thiogalactopyranoside. Recombinant CREB was purified using Ni-NTA agarose (Qiagen, Valencia, CA) as described previously.[2]

As demonstrated in Supplementary Figure 4, strong, selective labeling of glycosylated CREB was observed upon treatment with both Y289L GalT and analogue **1**. With larger quantities of protein, a faint background signal was observed, which was presumably due to the

non-specific interaction of aminooxy biotin with the protein. Importantly, the background signal was readily diagnosed using control reactions in the absence of enzyme or analogue **1**. In the case of *E. coli* CREB, for example, a weak background signal was observed over time, but no selective enhancement of signal was seen in the presence of both enzyme and **1**, indicating that bacterially expressed CREB was not GlcNAc glycosylated.

Labeling of α-crystallin. Bovine lens α-crystallin (a mixture of A and B chains) was resolved by SDS-PAGE electrophoresis and Coomassie-stained with standards in order to quantify the amount of A chain in the mixture. For reactions, 8.7 μg of α-crystallin (6.5 μg of A chain) in 20 mM HEPES pH 7.9 was added to 50 mM MOPS pH 6.45 containing 5 mM $MnCl_2$ and 0.25 mU/μL alkaline phosphatase. Analogue **1** and Y289L GalT were added to final concentrations of 1 mM and 100 ng/μL, respectively. Reactions were incubated at 4 °C for 18 h and then diluted 5-fold with PBS pH 6.7, protease inhibitors, and aminooxy biotin (6.5 mM final concentration). Biotinylation reactions were incubated with gentle shaking at 25 °C for 12 h. The molar ratio of biotin to α-crystallin was adjusted to minimize background signal, while maintaining reactivity over a reasonable time period. We found that a 4000:1 molar ratio worked successfully for these purposes. After biotinylation, reactions were aliquoted for analysis and subsequently boiled in SDS-PAGE loading dye. Proteins were resolved by 15% SDS-PAGE, transferred to nitrocellulose, and probed with streptavidin-HRP or stained with Coomassie Brilliant Blue (Supplementary Figure 5). Blotting with streptavidin-HRP was performed as described above and produced a strong signal within 30 min. In contrast, tritium labeling required 8 days to obtain a moderate signal. The difference in time corresponds to ~380-fold improvement in detection sensitivity.

UDP-[^{3}H]galactose labeling of α-crystallin. ^{3}H-labeling was performed essentially as described.[10,11] Briefly, 8.7 μg of α-crystallin (6.5 μg of A chain) in 20 mM HEPES pH 7.9 was added to 10 mM HEPES pH 7.9 containing 5 mM $MnCl_2$ and protease inhibitors. UDP-[^{3}H]-galactose was added to a final concentration of 0.03 μCi/μL, and the reaction was initiated with the addition of 25 mU autogalactosylated bovine β1,4-galactosyltransferase.[11] Reactions were incubated at 37 °C for 1 h 15 min. Reactions were subsequently aliquoted for analysis and stopped by boiling with SDS-PAGE loading dye. Proteins were resolved by 15% SDS-PAGE, stained with Coomassie Brilliant Blue, incubated with Amplify reagent, and dried for subsequent exposure to Hyperfilm MP at –80 °C.

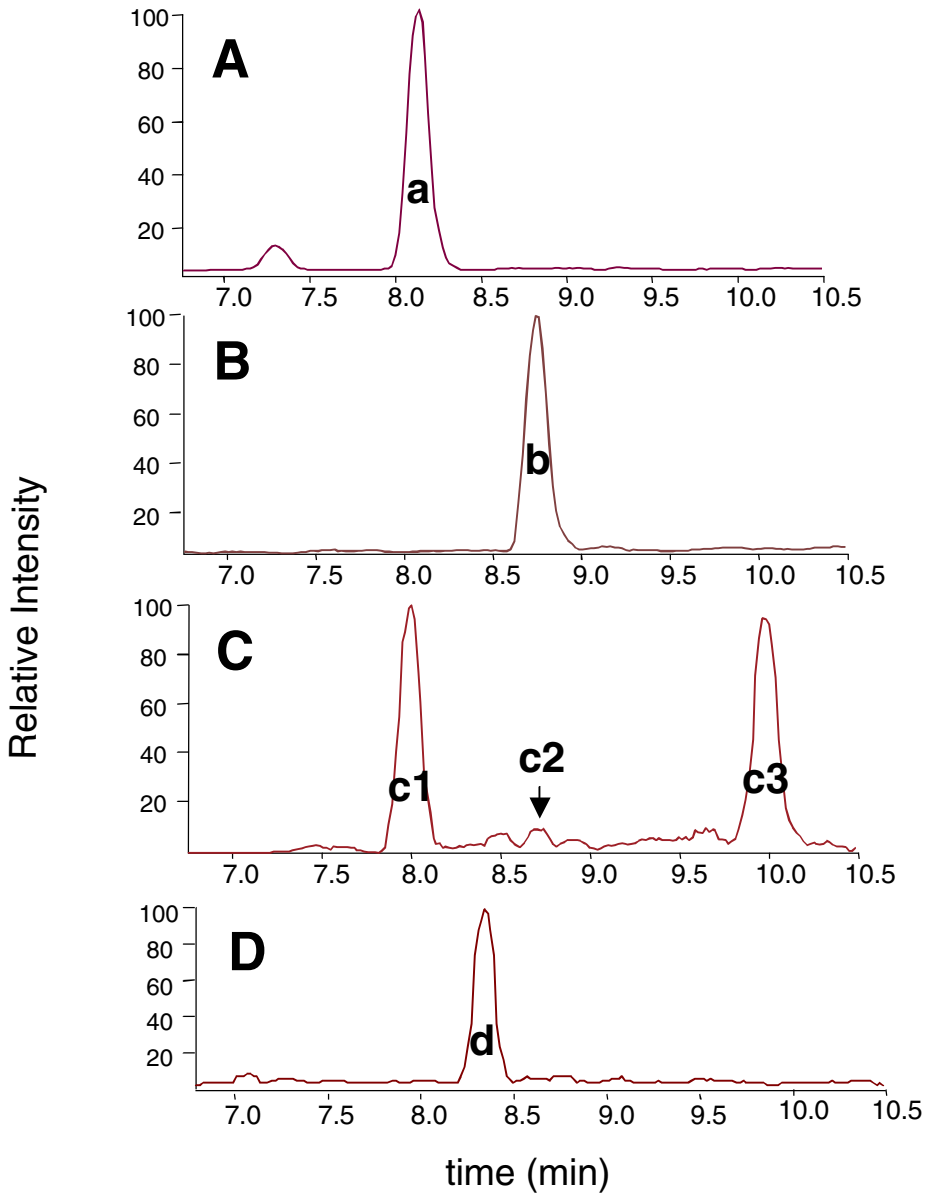
Western blotting of α-crystallin using antibodies RL-2 and CTD110.6. α-Crystallin, and appropriate positive and negative controls were resolved by 15% SDS-PAGE. All Western blotting steps were performed at RT unless otherwise noted. Western blotting with the RL-2 antibody was performed according to reported methods[12] with minor changes suggested by the manufacturer to reduce background noise. α-Crystallin and controls were electrophoretically transferred to nitrocellulose blots, and the blots were blocked for 1 h in 5% BSA in high salt (250 mM) TBS-1% tween-20 (hsTBS-T). RL-2 antibody, at a concentration of 1:2000, was subsequently added in blocking buffer and blots were incubated for 1.5-2 h. Blots were then rinsed with hsTBST and washed 6X5 min. Secondary goat anti-mouse IgG antibody was applied at a concentration of 1:10,000 in hsTBS-T containing 1% BSA. After 1 h, blots were rinsed and washed as described before chemiluminescence detection on film (Supplementary Figure 6A). Western blotting with the CTD110.6 antibody was performed according to manufacturer's

recommendations. Briefly, α-crystallin and controls were transferred to PVDF and washed 2X 15 min with TBS-0.1% tween-20 (TBST). Blots were blocked in TBST containing 3% BSA for 1 h, rinsed 2X with TBST, and probed with CTD110.6 (1:2500) in blocking buffer for 1 h. Blots were then rinsed 2X with TBST and washed 2X5 min. Secondary goat anti-mouse IgM antibody was applied at a concentration of 1:10,000 in blocking buffer for 1 h, and blots were subsequently rinsed with TBST and washed 5X5 min before chemiluminescence detection on film (Supplementary Figure 6B).

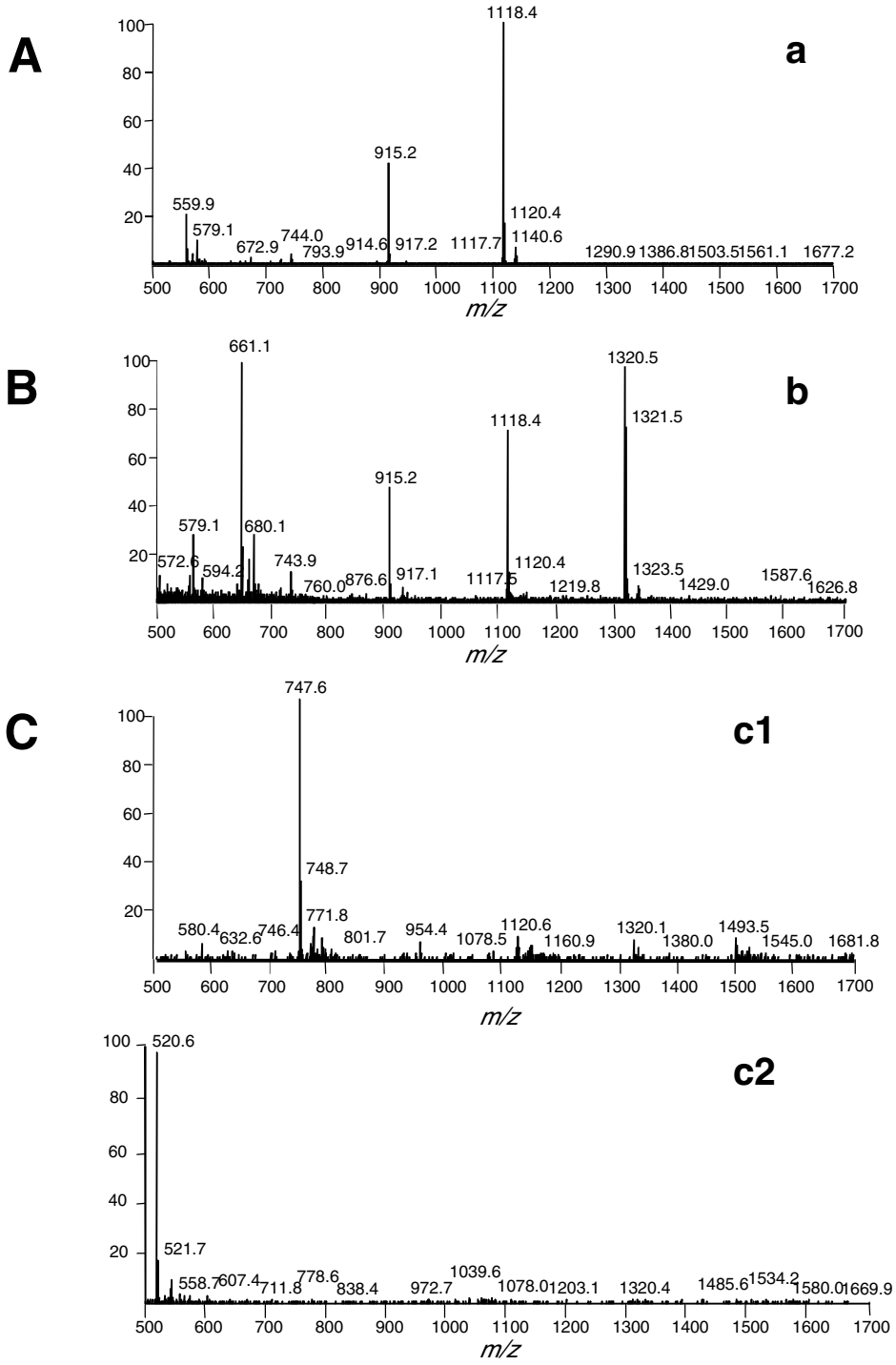
WGA lectin blotting of α-crystallin. WGA western blotting was performed essentially as described.[11,13] Briefly, α-crystallin and controls were resolved by 15% SDS-PAGE and electrophoretically transferred to nitrocellulose. Blots were blocked for 1 h in 3% periodate-treated BSA in PBS, rinsed 2X15 min with PBS-0.05% tween-20 (PBST), and probed for 2 h with WGA-HRP (1:8000 in PBST). Subsequently, blots were rinsed with PBST, washed 3X10 min, then 3X20 min before chemiluminescence detection on film (Supplementary Figure 7).

References:

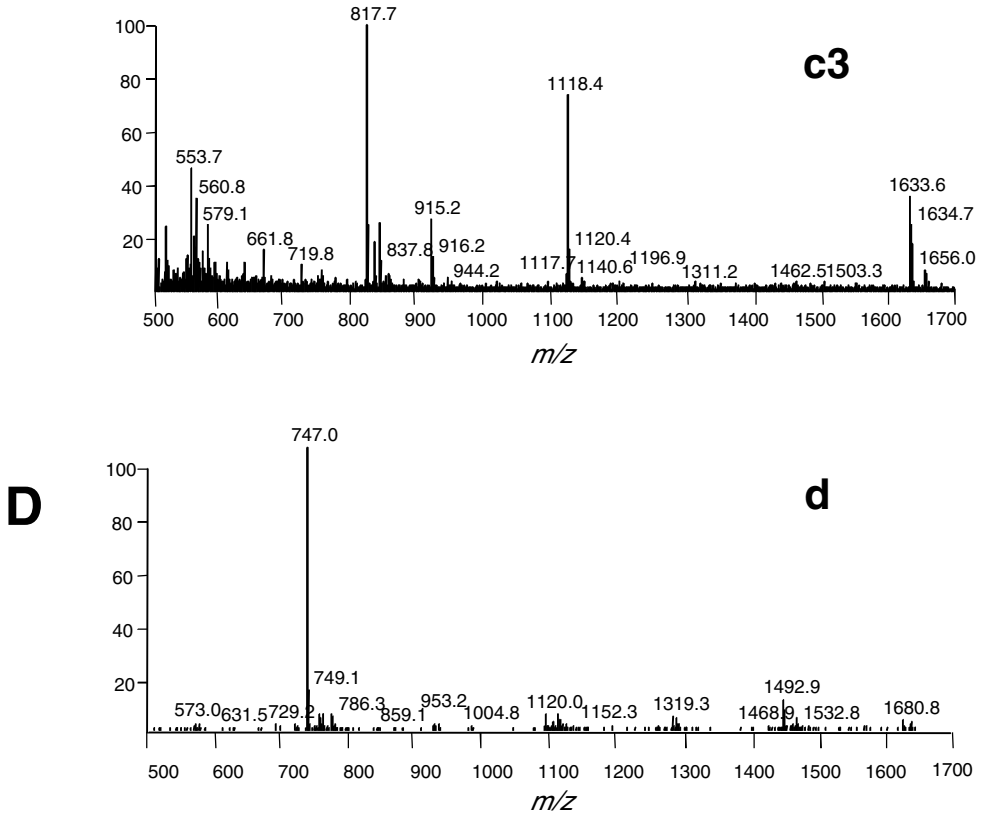
(1) Seitz, O.; Wong, C. H. *J. Am. Chem. Soc.* **1997**, *119*, 8766-8776.
(2) Lamarre-Vincent, N.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2003**, *125*, 6612-6613.
(3) Arts, J.; Herr, I.; Lansink, M.; Angel, P.; Kooistra, T. *Nucleic Acids Res.* **1997**, *25*, 311-317.
(4) Ramakrishnan, B.; Qasba, P. K. *J. Biol. Chem.* **2002**, *277*, 20833-20839.
(5) Ross, A. J.; Ivanova, I. A.; Ferguson, M. A. J.; Nikolaev, A. V. *J. Chem. Soc.-Perkin Trans. 1* **2001**, 72-81.
(6) (a) Ha, S.; Chang, E.; Lo, M. C.; Men, H.; Park, P.; Ge, M.; Walker, S. *J. Am. Chem. Soc.* **1999**, *121*, 8415-8426. (b) Sim, M. M.; Kondo, H.; Wong, C. H. *J. Am. Chem. Soc.* **1993**, *115*, 2260-2267.
(7) (a) Wittmann, V.; Wong, C. H. *J. Org. Chem.* **1997**, *62*, 2144-2147. (b) Hitchcock, S. A.; Eid, C. N.; Aikins, J. A.; Zia-Ebrahimi, M.; Blaszczak, L. C. *J. Am. Chem. Soc.* **1998**, *120*, 1916-1917.
(8) Unverzagt, C.; Kunz, H.; Paulson, J. C. *J. Am. Chem. Soc.* **1990**, *112*, 9308-9309.
(9) Glass, W. F.; Briggs, R. C.; Hnilica, L. S. *Anal. Biochem.* **1981**, *115*, 219-224.
(10) Roquemore, E. P.; Dell, A.; Morris, H. R.; Panico, M.; Reason, A. J.; Savoy, L. A.; Wistow, G. J.; Zigler, J. S.; Earles, B. J.; Hart, G. W. *J. Biol. Chem.* **1992**, *267*, 555-563.
(11) Roquemore, E. P.; Chou, T. Y.; Hart, G. W. *Methods Enzymol.* **1994**, *230*, 443-460.
(12) Konrad, R. J.; Janowski, K. M.; Kudlow, J. E. *Biochem. Biophys. Res. Commun.* **2000**, *267*, 26-32.
(13) Freeze, H. H. In *Current Protocols in Molecular Biology*; Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Eds.; John Wiley & Sons, Inc.: New York, 1999; Vol. 3, pp 17.7.1-17.7.8.



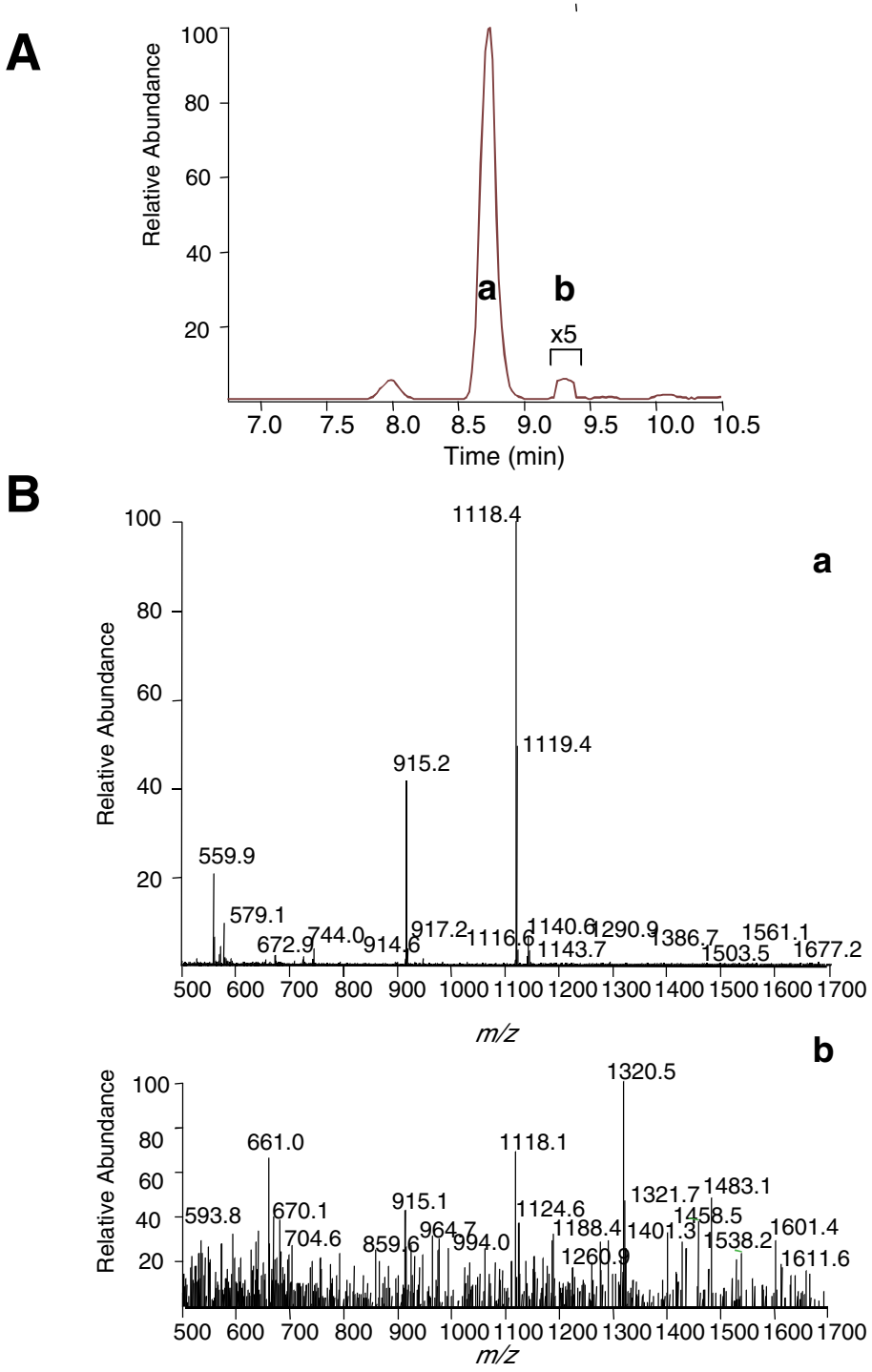
Supplementary Figure 1. Reverse phase LC-MS analysis of *O*-GlcNAc peptide labeling reactions at (A) time 0, (B) 6 h after the addition of **1** and Y289L GalT, (C) 8 h after aminooxy biotin addition. Trace D shows aminooxy biotin in the absence of **1**, Y289L GalT and *O*-GlcNAc peptide. A and B represent base peaks chromatograms, and C and D represent the extracted ion chromatograms within the mass range 1319.0-1321.0 *m/z* and 1633.0-1635.5 *m/z*. As shown in Supplementary Figure 2, peaks c1 and d represent the same biotin impurity. The slight difference in their retention times is due to minor differences in column equilibration time.



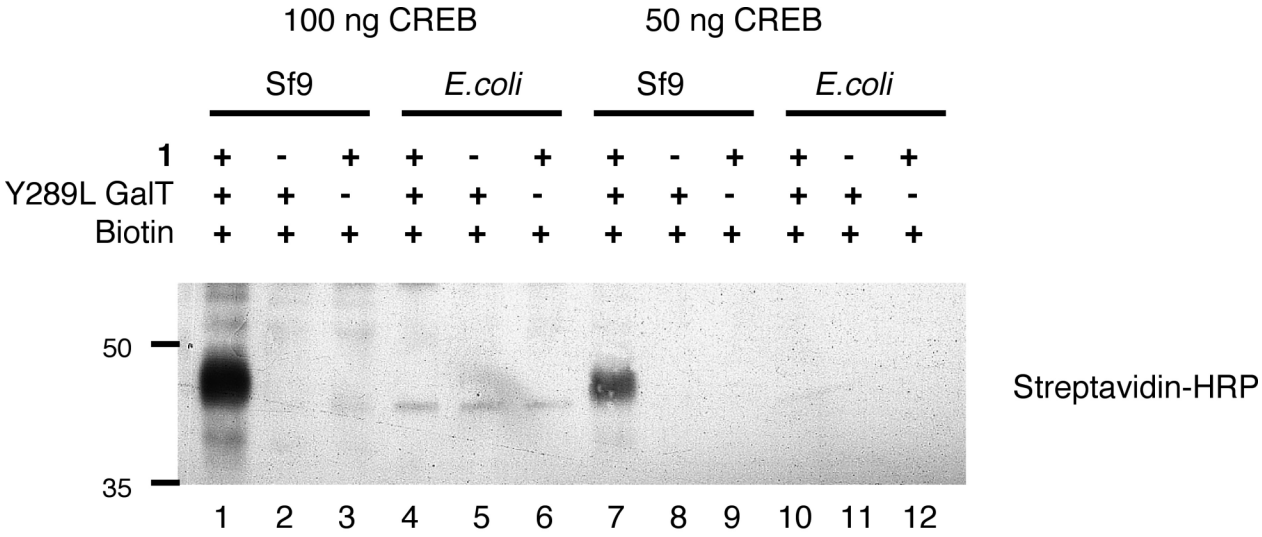
Supplementary Figure 2. Electrospray ionization mass spectra of the peaks in Supplementary Figure 1. (A) Spectrum of the peptide starting material (peak a), $[M_{GlcNAc}+H]^+$ =1118.4. The fragment ion at 915.2 *m/z* represents the deglycosylated peptide $[M+H]^+$, which was induced during ionization in the mass spectrometer. (B) Spectrum of the ketone product (peak b), $[M_{ketone\text{-}GlcNAc}+H]^+$ =1320.5. Ions at 1118.4, 915.2, and 661.1 *m/z* represent the *O*-GlcNAc glycosylated peptide, the deglycosylated peptide and the doubly charged species of the ketone labeled peptide, respectively. (C) Spectra of the biotin impurity (peak c1), peak c2, and the oxime product (peak c3)



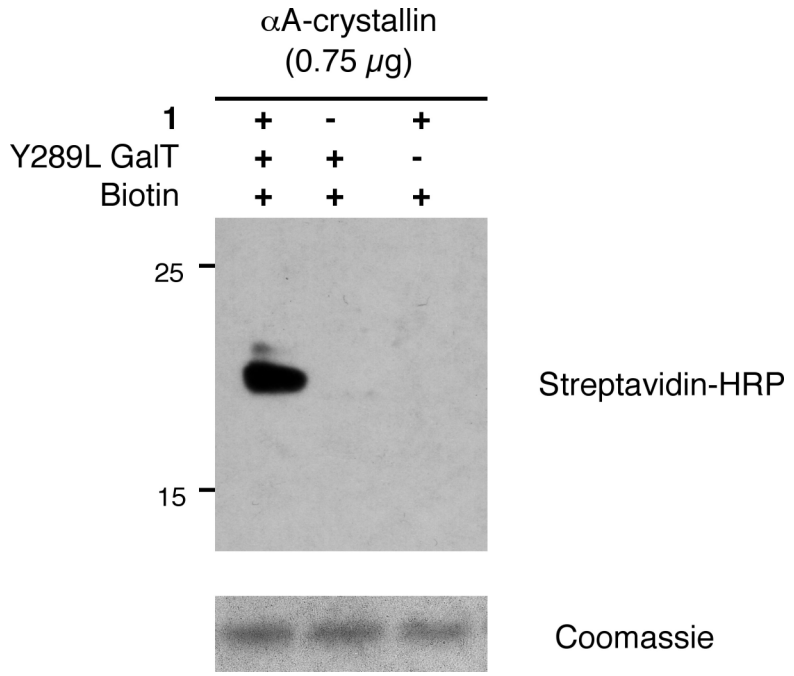
Supplementary Figure 2. *(continued)* The identity of the product was confirmed by ions 1633.6 and 817.7 *m/z*, which represent the singly and doubly charged species of the *O*-alkyl oxime product, respectively. The additional fragment ions at 1118.4 and 915.2 *m/z* correspond to the *O*-GlcNAc glycosylated and deglycosylated peptide, respectively. (D) Spectrum of the biotin impurity (peak d), obtained by incubating biotin in the absence of **1**, Y289L GalT and *O*-GlcNAc peptide. Note that the spectrum matches that of c1, indicating that these peaks arise from aminooxy biotin.



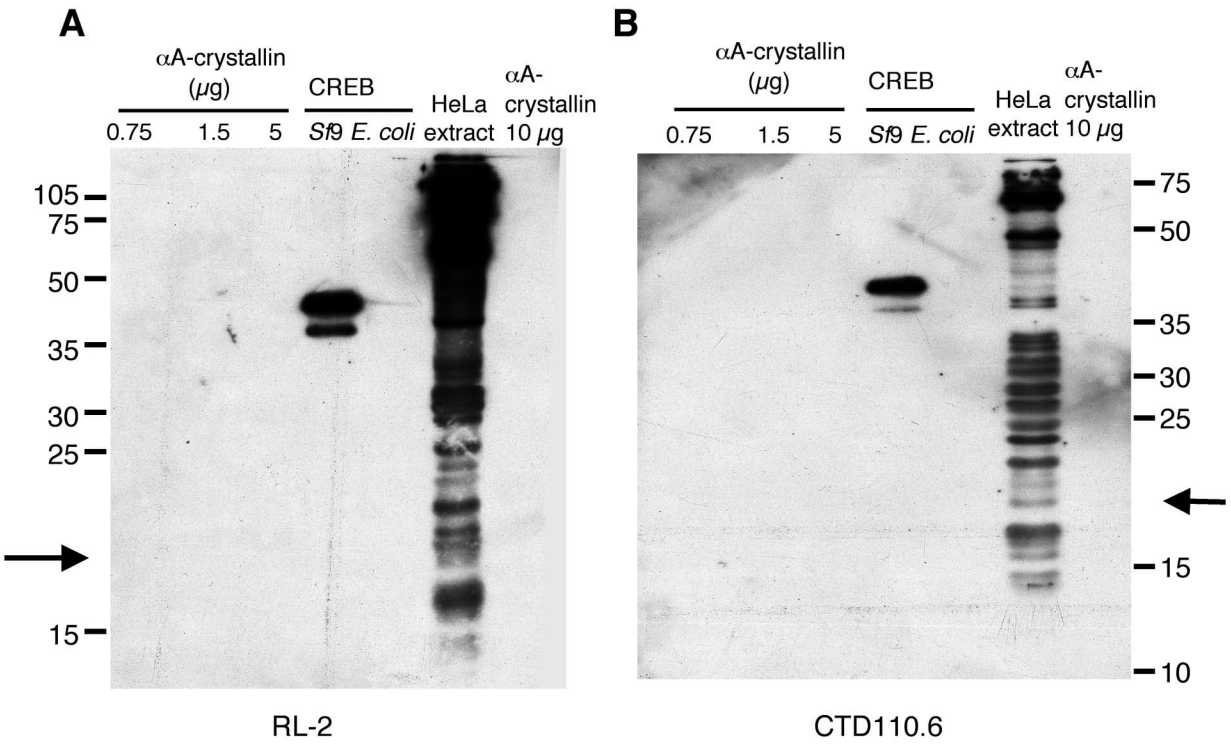
Supplementary Figure 3. (A) Reverse phase LC-MS analysis and accompanying mass spectra of the labeling reaction 12 h after the addition of **1** and wild-type GalT. Both the starting material (a) and ketone labeled peptide product peak (b) are visible in the base peak chromatogram. The latter peak intensity has been amplified 5-fold for clarity. (B) The EI mass spectra of peaks a and b confirm the identities of the *O*-GlcNAc glycosylated peptide, $[M_{GlcNAc}+H]^+ = 1118$ *m/z*, and the product, $[M_{ketone\text{-}GlcNAc}+H]^+ = 1320.635$ m/z and $[M_{ketone\text{-}GlcNAc}+2H]^{2+} = 661$ *m/z*, respectively.



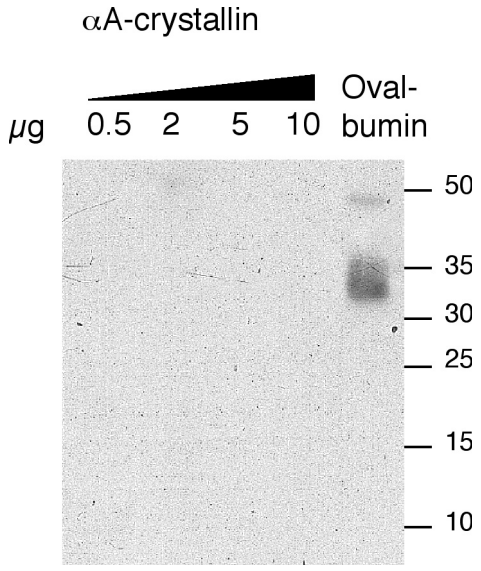
Supplementary Figure 4. Labeling of glycosylated CREB from Sf9 cells (lanes 1-3 and 7-9) or *E.coli* (lanes 4-6 and 10-12). Strong streptavidin-HRP signal is observed upon treatment with Y289L GalT and analogue **1** (lanes 1 and 7) relative to reactions lacking enzyme or **1** (lanes 2-3 and 8-9). In contrast, no selective enhancement of the signal is observed for the negative control, unglycosylated CREB from *E.coli*.



Supplementary Figure 5. Streptavidin-HRP blot of the α-crystallin labeling reactions and the accompanying Coomassie-stained gel of each reaction. Selective labeling of α-crystallin A chain was observed. In contrast, no appreciable labeling was observed in the control reactions lacking **1** or Y289L. Coomassie gel bands of similar intensity confirm the presence of comparable amounts of α-crystallin A chain. Faint labeling of the B chain was observed, consistent with reports that it is *O*-GlcNAc glycosylated.[10]



Supplementary Figure 6. Western blots of α-crystallin using the RL-2 antibody (A) and CTD110.6 antibody (B). Both antibodies effectively detected the *O*-GlcNAc present on the CREB positive control and HeLa nuclear lysates, while the negative control, unglycosylated CREB from *E. coli*, remained undetected. However, the antibodies failed to appreciably detect the *O*-GlcNAc present on α-crystallin, even when 10 μg of protein was used. The arrow marks the anticipated position of α-crystallin in the gel.



Supplementary Figure 7. Blotting of α-crystallin using WGA lectin. While WGA detected the *N*-linked terminal GlcNAc groups of the ovalbumin positive control, it could not detect the *O*-GlcNAc moiety on α-crystallin.